

AMENDMENT TO THE CLAIMS

1. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

a receiver operable to receive a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit and operable to receive a second sequence of sub-word unit labels representative of a second one of said at least two word alternatives output by said word recognition unit;

an aligner operable to align and to compare sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

a processor operable to process the aligned pairs of sub-word unit labels formed by said aligner, to determine an output sequence of sub-word unit labels representative of the at least two word alternatives.

2. (Previously Presented) An apparatus according to claim 1, wherein said processor is operable to determine said output sequence of sub-word unit labels by determining, for each aligned pair of sub-word unit labels, a sub-word unit label that is confusingly similar to the first and second sub-word unit labels of the aligned pair.

3. (Previously Presented) An apparatus according to claim 2, wherein said processor comprises:

a first comparator operable to compare, for each aligned pair, the first sequence sub-word unit label in the aligned pair with each of a plurality of sub-word unit labels taken from a set of predetermined sub-word unit labels, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit label and the respective sub-word unit labels of the set;

a second comparator operable to compare, for each aligned pair, the second sequence sub-word unit label in the aligned pair with each of said plurality of sub-word unit labels from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit label and the respective sub-word unit labels of the set;

a combiner operable to combine the comparison scores obtained when comparing the first and second sequence sub-word unit labels in the aligned pair with the same sub-word unit label from the set, to generate a plurality of combined comparison scores;

a third comparator operable to compare, for each aligned pair, the combined comparison scores generated by said combiner for the aligned pair; and

a determiner operable to determine, for each aligned pair of sub-word unit labels, a sub-word unit label representative of the sub-word unit labels in the aligned pair in dependence upon a comparison result output by said third comparator for the aligned pair.

4. (Previously Presented) An apparatus according to claim 3, wherein said first and second comparators are operable to compare the first sequence sub-word unit label and the second sequence sub-word unit label respectively with each of the sub-word unit labels in said set of sub-word unit labels.

5. (Previously Presented) An apparatus according to claim 3, wherein said first and second comparators are operable to provide comparison scores which are indicative of a probability of confusing the corresponding sub-word unit label taken from the set of predetermined sub-word unit labels as the sub-word unit label in the aligned pair.

6. (Previously Presented) An apparatus according to claim 5, wherein said combiner is operable to combine the comparison scores in order to multiply the probabilities of confusing the corresponding sub-word unit label taken from the set as the sub-word unit labels in the aligned pair.

7. (Previously Presented) An apparatus according to claim 6, wherein each of said sub-word unit labels in said set of predetermined sub-word unit labels has a predetermined probability of occurring within a sequence of sub-word unit labels and wherein said combiner is operable to weight each of said combined comparison scores in dependence upon the respective probability of occurrence for the sub-word unit label of the set used to generate the combined comparison score.

8. (Previously Presented) An apparatus according to claim 7, wherein said combiner is operable to combine said comparison scores by calculating:

$$P(d_i^1|p_r)P(d_j^2|p_r)P(p_r)$$

where d_i^1 and d_j^2 are an aligned pair of first and second sequence sub-word unit labels respectively, $P(d_i^1|p_r)$ is the comparison score output by said first comparator and is representative of the probability of confusing set sub-word unit label p_r as first sequence sub-word unit label d_i^1 ; $P(d_j^2|p_r)$ is the comparison score output by said second comparator and is representative of the probability of confusing set sub-word unit label p_r as second sequence sub-word unit label d_j^2 ; and $P(p_r)$ is a weight which represents the probability of set sub-word unit label p_r occurring in a sequence of sub-word unit labels.

9. (Previously Presented) An apparatus according to claim 8, wherein said third comparator is operable to identify the set sub-word unit label p_r which gives the maximum combined comparison score and wherein said determiner is operable to determine said sub-word unit label representative of the sub-word unit labels in the aligned pair as being the sub-word unit label which provides the maximum combined comparison score.

10. (Currently Amended) An apparatus according to claim 6, wherein said ~~compare~~ comparison scores represent log probabilities and wherein said combiner is operable to multiply said probabilities by adding the respective comparison scores.

11. (Previously Presented) An apparatus according to claim 3, wherein each of the sub-word unit labels in said first and second sequences of sub-word unit labels belong to said set of predetermined sub-word unit labels and wherein said first and second comparators are operable to provide said comparison scores using predetermined data which relate the sub-word unit labels in said set to each other.

12. (Previously Presented) An apparatus according to claim 11, wherein said predetermined data comprises, for each sub-word unit label in the set of sub-word unit labels, a probability for confusing that sub-word unit label with each of the other sub-word unit labels in the set of sub-word unit labels.

13. (Previously Presented) An apparatus according to claim 1, wherein said aligner comprises a dynamic programmer operable to align said first and second sequences of sub-word unit labels using a dynamic programming technique.

14. (Previously Presented) An apparatus according to claim 13, wherein said dynamic programmer is operable to determine an optimum alignment between said first and second sequences of sub-word unit labels.

15. (Previously Presented) An apparatus according to claim 1, wherein each of said sub-word unit labels represents a phoneme.

16. (Previously Presented) An apparatus according to claim 1, wherein said receiver is operable to receive a third sequence of sub-word unit labels representative of a third word alternative output by said word recognition unit and wherein said aligner is operable to simultaneously align and compare the sub-word unit labels of the first, second and third sequences of sub-word unit labels.

17. (Previously Presented) An apparatus according to claim 1, wherein said receiver is operable to receive a third sequence of sub-word unit labels representative of a third word alternative output by said word recognition unit and wherein said aligner is operable to align and compare two sequences of sub-word unit labels at a time.

18. (Previously Presented) An apparatus according to claim 1, further comprising a word to sub-word unit label dictionary which is operable to receive the at least two word alternatives output by said word recognition unit and to generate therefrom said first and second sequences of sub-word unit labels.

19. (Previously Presented) An apparatus according to claim 1, further comprising an annotator operable to annotate a data file using the output sequence of sub-word unit labels output determined by said processor.

20. (Previously Presented) An apparatus according to claim 19, wherein said annotator is operable to annotate said data file using said output sequence of sub-word unit labels and said at least two word alternatives output by said word recognition unit.

21. (Previously Presented) An apparatus according to claim 20, wherein said output sequence of sub-word unit labels and said at least two word alternatives are combined to form annotation data for the data file.

22. (Original) An apparatus according to claim 19, wherein said data file is one of: an audio data file, a video data file, an image data file or a text data file.

23. (Original) An apparatus according to claim 1, wherein said word recognition unit comprises a speech recognition system.

24. (Original) An apparatus according to claim 1, wherein said word recognition unit comprises a handwriting recognition system.

25. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two words, the apparatus comprising:

a receiver operable to receive a first sequence of sub-word unit labels representative of a first word and a second sequence of sub-word unit labels representative of a second word;

an aligner operable to align sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

a processor operable to process the aligned pairs of sub-word unit labels formed by said aligner to determine an output sequence of sub-word unit labels representative of the first and second sequences of sub-word unit labels by determining, for each aligned pair of sub-word unit labels, a sub-word unit label that is confusingly similar to the sub-word unit labels of the aligned pair.

26. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

a receiver operable to receive the word alternatives output by the word recognition unit;

a generator operable to generate, for each received word alternative, a sequence of sub-word unit labels representative of the word alternative;

an aligner operable to align and compare the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number aligned groups of sub-word unit labels; and

a processor operable to process the aligned groups of sub-word unit labels identified by said aligner, to determine an output sequence of sub-word unit labels representative of the received word alternatives.

27. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word

recognition unit in response to a common input word to be recognised, the apparatus comprising:

a receiver operable to receive the at least two word alternatives output by the word recognition unit;

a generator operable to generate, for each received word alternative, a sequence of sub-word unit labels representative of the received word alternative;

an aligner operable to align and compare the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number aligned groups of sub-word unit labels; and

a processor operable to process the aligned groups of sub-word unit labels identified by said aligner, to determine an output sequence of sub-word unit labels representative of the at least two word alternatives by determining, for each aligned group of sub-word unit labels, a sub-word unit label that is confusingly similar to the sub-word unit labels of the group.

28. (Currently Amended) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

a receiver operable to receive a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit and operable to receive a second sequence of sub-word unit labels

representative of a second one of said at least two word alternatives output by said word recognition unit;

an aligner operable to align and compare sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels;

a first comparator operable to compare,[[,]] for each aligned pair, the first sequence sub-word unit label in the aligned pair with each of a plurality of sub-word unit labels taken from a set of predetermined sub-word unit labels, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit label and the respective sub-word unit labels of the set;

a second comparator operable to compare, for each aligned pair, the second sequence sub-word unit label in the aligned pair with each of said plurality of sub-word unit labels from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit label and the respective sub-word unit labels of the set;

a combiner operable to combine the comparison scores obtained when comparing the first and second sequence sub-word unit labels in the aligned pair with the same sub-word unit label from the set, to generate a plurality of combined comparison scores;

a third comparator operable to compare,[[,]] for each aligned pair, the combined comparison scores generated by said combiner for the aligned pair; and

a determiner operable to determine, for each aligned pair of sub-word unit labels, a sub-word unit label representative of the sub-word unit labels in the aligned pair in

dependence upon a comparison result output by said third comparator for the aligned pair.

29. (Previously Presented) A method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit;

receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternatives output by said word recognition unit;

aligning and comparing sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

processing the aligned pairs of sub-word unit labels formed in said aligning step, to determine an output sequence of sub-word unit labels representative of the at least two word alternatives.

30. (Previously Presented) A method according to claim 29, wherein said processing step determines said output sequence of sub-word unit labels by determining, for each aligned pair of sub-word unit labels, a sub-word unit label that is confusingly similar to the first and second sub-word unit labels of the aligned pair.

31. (Previously Presented) A method according to claim 29, wherein said processing step comprises:

a first comparing step of comparing, for each aligned pair, the first sequence sub-word unit label in the aligned pair with each of a plurality of sub-word unit labels taken from a set of predetermined sub-word unit labels, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit label and the respective sub-word unit labels of the set;

a second comparing step of comparing, for each aligned pair, the second sequence sub-word unit label in the aligned pair with each of said plurality of sub-word unit labels from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit label and the respective sub-word unit labels of the set;

combining the comparison scores obtained when comparing the first and second sequence sub-word unit labels in the aligned pair with the same sub-word unit label from the set, to generate a plurality of combined comparison scores;

a third comparing step of comparing, for each aligned pair, the combined comparison scores generated in said combining step for the aligned pair; and

determining, for each aligned pair, a sub-word unit label representative of the sub-word unit labels in the aligned pair in dependence upon a comparison result output from said third comparing step for the aligned pair.

32. (Previously Presented) A method according to claim 31, wherein said first and second comparing steps compare the first sequence sub-word unit label and

the second sequence sub-word unit label respectively with each of the sub-word unit labels in said set of sub-word unit labels.

33. (Previously Presented) A method according to claim 31, wherein said first and second comparing steps provide comparison scores which are indicative of a probability of confusing the corresponding sub-word unit label taken from the set of predetermined sub-word unit labels as the sub-word unit label in the aligned pair.

34. (Previously Presented) A method according to claim 33, wherein said combining step combines the comparison scores in order to multiply the probabilities of confusing the corresponding sub-word unit label taken from the set as the sub-word unit labels in the aligned pair.

35. (Previously Presented) A method according to claim 34, wherein each of said sub-word unit labels in said set of predetermined sub-word unit labels has a predetermined probability of occurring within a sequence of sub-word unit labels and wherein said combining step weights each of said combined comparison scores in dependence upon the respective probability of occurrence for the sub-word unit label of the set used to generate the combined comparison score.

36. (Previously Presented) A method according to claim 35, wherein said combining step combines said comparison scores by calculating:

$$P(d_i^1|p_r)P(d_j^2|p_r)P(p_r)$$

where d_i^1 and d_j^2 are an aligned pair of first and second sequence sub-word unit labels respectively; $P(d_i^1|p_r)$ is the comparison score output by said first comparing step and is representative of the probability of confusing set sub-word unit label p_r as first sequence sub-word unit label d_i^1 ; $P(d_j^2|p_r)$ is the comparison score output by said second comparing step and is representative of the probability of confusing set sub-word unit label p_r as second sequence sub-word unit label d_j^2 ; and $P(p_r)$ is a weight which represents the probability of set sub-word unit label p_r occurring in a sequence of sub-word unit labels.

37. (Previously Presented) A method according to claim 36, wherein said third comparing step identifies the set sub-word unit label p_r which gives the maximum combined comparison score and wherein said determining step determines said sub-word unit label representative of the sub-word unit labels in the aligned pair as being the sub-word unit label which provides the maximum combined comparison score.

38. (Original) A method according to claim 34, wherein said comparison scores represent log probabilities and wherein said combining step multiplies said probabilities by adding the respective comparison scores.

39. (Previously Presented) A method according to claim 31, wherein each of the sub-word unit labels in said first and second sequences of sub-word units belong to said set of predetermined sub-word unit labels and wherein said first and second

comparing steps provide said comparison scores using predetermined data which relate the sub-word unit labels in said set to each other.

40. (Previously Presented) A method according to claim 39, wherein said predetermined data comprises, for each sub-word unit label in the set of sub-word unit labels, a probability for confusing that sub-word unit label with each of the other sub-word unit labels in the set of sub-word unit labels.

41. (Previously Presented) A method according to claim 29, wherein said aligning and comparing step uses a dynamic programming technique to align said first and second sequences of sub-word unit labels.

42. (Previously Presented) A method according to claim 41, wherein said dynamic programming technique determines an optimum alignment between said first and second sequences of sub-word unit labels.

43. (Previously Presented) A method according to claim 29, wherein each of said sub-word unit labels represents a phoneme.

44. (Previously Presented) A method according to claim 29, further comprising receiving a third sequence of sub-word unit labels representative of a third word alternative output by said word recognition unit and wherein said aligning and

comparing step simultaneously aligns and compares the sub-word unit labels of the first, second and third sequences of sub-word unit labels.

45. (Previously Presented) A method according to claim 29, further comprising receiving a third sequence of sub-word unit labels representative of a third word alternative output by said recognition unit and wherein said aligning and comparing step aligns and compares two sequences of sub-word unit labels at a time.

46. (Previously Presented) A method according to claim 29, further comprising the step of using a word to sub-word unit label dictionary to convert the words output by said word recognition unit into said sequences of sub-word unit labels.

47. (Previously Presented) A method according to claim 29, further comprising the step of annotating a data file using the sub-word unit labels determined by said processing step.

48. (Previously Presented) A method according to claim 47, wherein said annotating step annotates said data file using said output sequence of sub-word unit labels and said word alternatives output by said word recognition unit.

49. (Previously Presented) A method according to claim 48, wherein said output sequence of sub-word unit labels and said word alternatives are combined to

form annotation data for the data file.

50. (Original) A method according to claim 47, wherein said data file is one of: an audio data file, a video data file, an image data file, or a text data file.

51. (Original) A method according to claim 29, wherein said word recognition unit comprises a speech recognition system.

52. (Original) A method apparatus according to claim 29, wherein said word recognition unit comprises a handwriting recognition system.

53. (Previously Presented) A method of determining a sequence of sub-word unit labels representative of at least two words, the method comprising:

- receiving a first sequence of sub-word unit labels representative of a first word;
- receiving a second sequence of sub-word unit labels representative of a second word;
- aligning sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and
- processing the aligned pairs of sub-word unit labels formed in said aligning step, to determine an output sequence of sub-word unit labels representative of the first and second sequences of sub-word unit labels by determining, for each aligned pair of sub-word

unit labels, a sub-word unit label that is confusingly similar to the sub-word unit labels of the aligned pair.

54. (Previously Presented) A method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving the word alternative output by the word recognition unit;

generating, for each received word alternative, a sequence of sub-word unit labels representative of the word alternative;

aligning and comparing the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number of aligned groups of sub-word unit labels; and

processing the aligned groups of sub-word unit labels identified in said aligning and comparing step to determine an output sequence of sub-word unit labels representative of the received word alternatives.

55. (Previously Presented) A method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving the word alternatives output by the word recognition unit;

generating, for each received word alternative, a sequence of sub-word unit

labels representative of the word alternative;

aligning and comparing the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number of aligned groups of sub-word unit labels; and

processing the aligned groups of sub-word labels identified in said aligning and comparing step to determine an output sequence of sub-word unit labels representative of the at least two word alternatives by determining, for each aligned group, a sub-word unit label that is confusingly similar to the sub-word unit labels of the group.

56. (Previously Presented) A method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the method comprising:

receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit;

receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternative output by said word recognition unit;

aligning and comparing sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels;

a first comparing step of comparing, for each aligned pair, the first sequence sub-word unit label in the aligned pair with each of a plurality of sub-word unit labels taken from a set of predetermined sub-word unit labels, to provide a corresponding plurality of

comparison scores representative of the similarities between the first sequence sub-word unit label and the respective sub-word unit labels of the set;

a second comparing step of comparing, for each aligned pair, the second sequence sub-word unit label in the aligned pair with each of said plurality of sub-word unit labels from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit label and the respective sub-word unit labels of the set;

combining the comparison scores obtained when comparing the first and second sequence sub-word unit labels in the aligned pair with the same sub-word unit label from the set, to generate a plurality of combined comparison scores;

a third comparing step of comparing, for each aligned pair, the combined comparison scores generated in said combining step for the aligned pair; and

determining, for each aligned pair, a sub-word unit representative of the sub-word unit labels in the aligned pair in dependence upon a comparison result output from said third comparing step for the aligned pair.

57. (Previously Presented) A computer readable medium storing computer executable process steps to perform a method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the process steps comprising:

steps for receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit;

steps for receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternatives output by the word recognition unit;

steps for aligning and comparing sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

steps for processing the aligned pairs of sub-word unit labels formed in said aligning and comparing step to determine an output sequence of sub-word unit labels representative of the at least two word alternatives.

58. (Previously Presented) A computer executable program stored on a computer readable medium, the computer executable program for controlling a processor to perform a method of determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the program comprising:

code for receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by the word recognition unit;

code for receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternative output by the word recognition unit;

code for aligning and comparing sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

code for processing the aligned pairs of sub-word unit labels formed by said

aligning and comparing code to determine an output sequence of sub-word unit labels representative of the at least two word alternatives.

59. (Previously Presented) An apparatus according to claim 1, wherein said processor is operable to process the aligned pairs of sub-word unit labels formed by said aligner to determine an output sequence of sub-word unit labels that is different from the first sequence of sub-word unit labels and different from the second sequence of sub-word unit labels and which is representative of the at least two word alternatives.

60. (Previously Presented) A method according to claim 29, wherein said processing step processes the aligned pairs of sub-word unit labels formed in said aligning step, to determine an output sequence of sub-word unit labels which is different from the first sequence of sub-word unit labels and different from the second sequence of sub-word unit labels and which is representative of the at least two word alternatives.

61. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit and for receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternatives output by said word recognition unit;

means for aligning and for comparing sub-word unit labels of the first sequence with sub-word labels of the second sequence to form a number of aligned pairs of sub-word unit labels; and

means for processing aligned pairs of sub-word unit labels formed by said aligning means to determine an output sequence of sub-word unit labels representative of the least two word alternatives.

62. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two words, the apparatus comprising:

means for receiving a first sequence of sub-word unit labels representative of a first word and for receiving a second sequence of sub-word unit labels representative of a second word;

means for aligning sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number aligned pairs of sub-word unit labels; and

means for processing the aligned pairs of sub-word unit labels formed by said aligning means to determine an output sequence of sub-word unit labels representative of the first and second sequence of sub-word unit labels by determining, for each aligned pair of sub-word unit labels, a sub-word unit label that is confusingly similar to the sub-word unit labels of the aligned pair.

63. (Currently Amended) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word

recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving the word alternatives output by the word recognition unit;

means for generating for each received word alternative, a sequence of sub-word unit labels representative of the word alternative;

means for aligning and comparing the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number aligned groups of sub-word unit labels; and

means for processing the aligned groups of sub-word unit labels identified by said aligning means, to determine an output sequence of sub-word unit labels representative of the received word alternatives.

64. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving the at least two word alternatives output by the word recognition unit;

means for generating, for each received word alternative, a sequence of sub-word unit labels representative of the received word alternative;

means for aligning and comparing the sub-word unit labels of each generated sequence of sub-word unit labels to identify a number aligned groups of sub-

word unit labels; and

means for processing the aligned groups of sub-word unit labels identified by said aligning means to determine an output sequence of sub-word unit labels representative of the at least two word alternatives by determining, for each aligned group of sub-word unit labels, a sub-word unit label that is confusingly similar to the sub-word unit labels of the group.

65. (Previously Presented) An apparatus for determining a sequence of sub-word unit labels representative of at least two word alternatives output by a word recognition unit in response to a common input word to be recognised, the apparatus comprising:

means for receiving a first sequence of sub-word unit labels representative of a first one of said at least two word alternatives output by said word recognition unit and for receiving a second sequence of sub-word unit labels representative of a second one of said at least two word alternatives output by said word recognition unit;

means for aligning and for comparing sub-word unit labels of the first sequence with sub-word unit labels of the second sequence to form a number of aligned pairs of sub-word unit labels;

first comparing means for comparing, for each aligned pair, the first sequence sub-word unit label in the aligned pair with each of a plurality of sub-word unit labels taken from a set of predetermined sub-word unit labels, to provide a corresponding plurality of comparison scores representative of the similarities between the first sequence sub-word unit label and the respective sub-word unit labels of the set;

second comparing means for comparing, for each aligned pair, the second sequence sub-word unit label in the aligned pair with each of said plurality of sub-word unit labels from the set, to provide a further corresponding plurality of comparison scores representative of the similarities between said second sequence sub-word unit label and their respective sub-word unit labels of the set;

means for combining the comparison scores obtained when comparing the first and second sequence sub-word unit labels in the aligned pair with the same sub-word unit labels from the set, to generate a plurality of combined comparison scores;

third comparing means for comparing, for each aligned pair, the combined comparison scores generated by said combining means for the aligned pair; and

means for determining, for each aligned pair of sub-word unit labels, a sub-word unit label representative of the sub-word unit labels in the aligned pair in dependence upon a comparison result output by said third comparing means for the aligned pair.